



(11) Publication number : **0 602 791 A2**

(12) **EUROPEAN PATENT APPLICATION**

(21) Application number : 93308706.6

(51) Int. Cl.<sup>5</sup> : **G06F 15/16, G06F 9/445**

(22) Date of filing : 01.11.93

(30) Priority : 30.10.92 US 969596

(43) Date of publication of application :  
22.06.94 Bulletin 94/25

(84) Designated Contracting States :  
AT BE CH DE ES FR GB IT LI NL SE

(71) Applicant : International Business Machines Corporation  
Old Orchard Road  
Armonk, N.Y. 10504 (US)

(72) Inventor : Hardell, William Rudolph, Jr.  
8107A Baywood Drive  
Austin, Texas 78759 (US)  
Inventor : Henson, James Dodd, Jr.  
8405 Maine Drive  
Austin, Texas 78758 (US)  
Inventor : Mitchell, Oscar Reid  
600 Willowood Lane  
Pflugerville, Texas 78660 (US)

(74) Representative : Moss, Robert Douglas  
IBM United Kingdom Limited  
Intellectual Property Department  
Hursley Park  
Winchester Hampshire SO21 2JN (GB)

(54) Apparatus and method for booting a multiple processor system having a global/local memory architecture.

(57) An architecture and method for booting a multi-processor system having processor local memory and shared global memory, with shared global memory access managed by an atomic memory access controller and cache coherence managed by software. Reset circuits are used to synchronize to a master clock a commonly distributed start signal and processor individualized restart sequences, which reset circuit signals are distributed to reset both local and global memory. Global memory testing is assigned to a processor based upon its rate status in completing an internal test sequence. The systems and methods are particularly suited to booting a group of multiple but relatively independent processors. Furthermore, the practice of the invention facilitates booting of such system when one or more of the processors have been disconnected or failed.

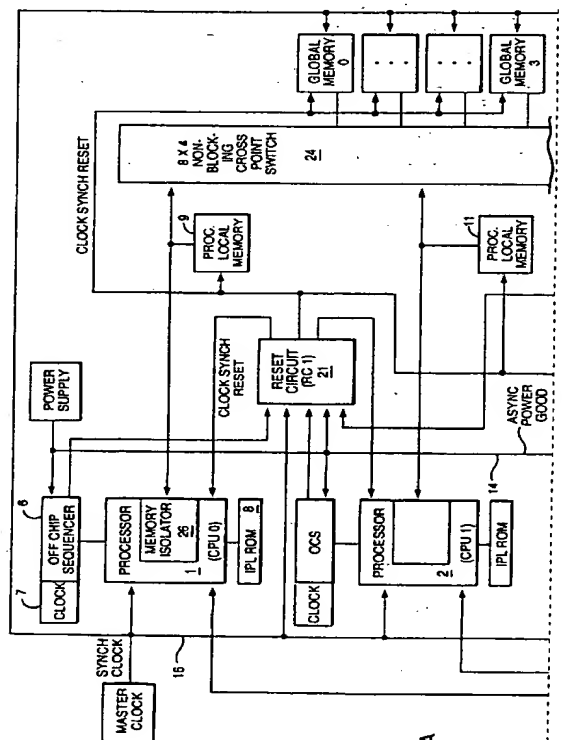


FIG. 1A

EP 0 602 791 A2

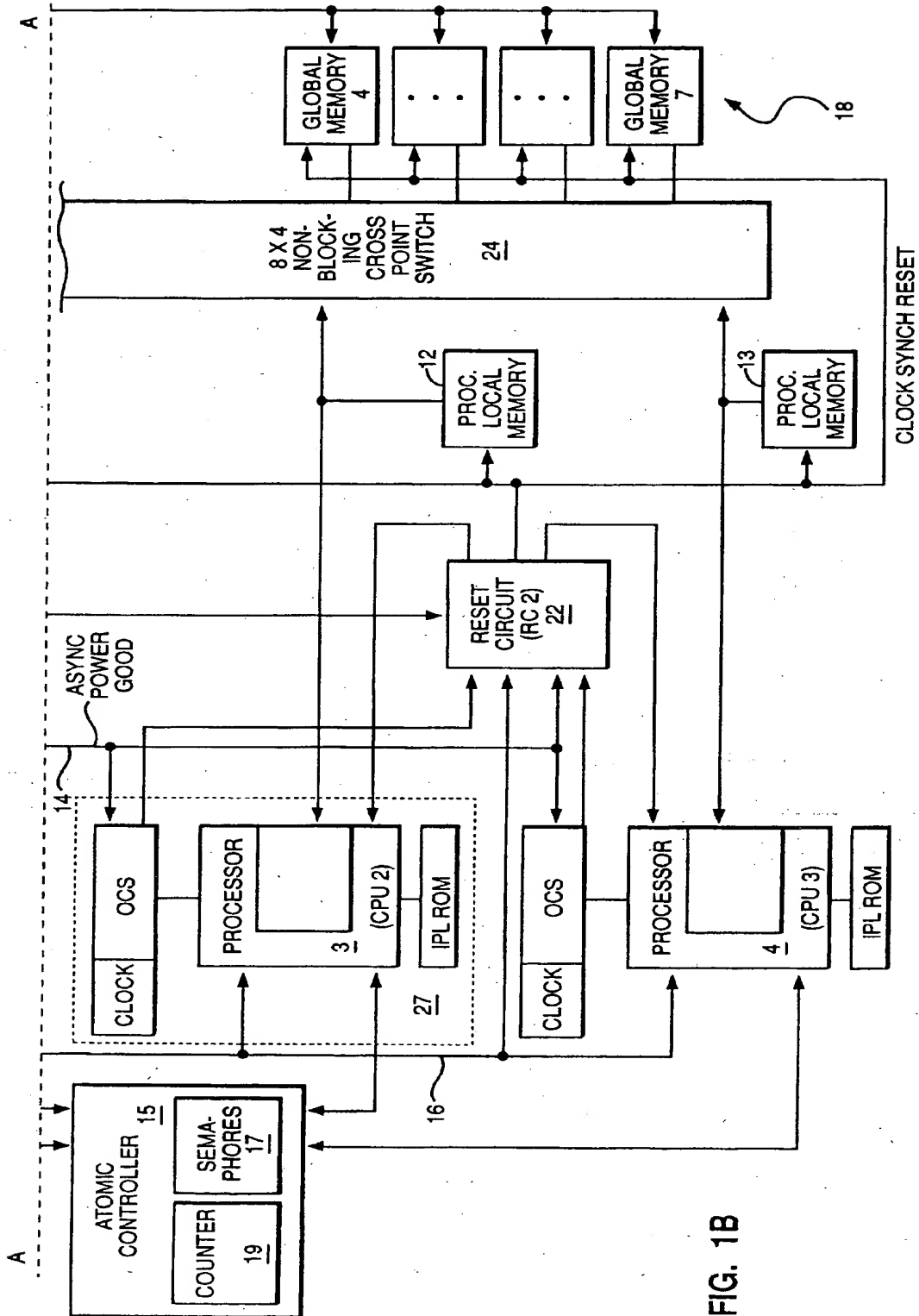


FIG. 1B

### Cross-Reference to Related Applications

The present invention is related to co-pending European Patent Application No. 93308323.0 (IBM Docket AT9-92-147), having title "Apparatus and Method for Steering Spare Bits in a Multiple Processor System", and having common inventorship and applicant.

### Background of the Invention

The present invention relates generally to multiple processor computer systems. More particularly, the invention is directed to systems and methods for booting/starting/restarting/resetting a multiple processor system characterized by the presence of a shared global memory and a multiplicity of relatively independently operable processors having individualized resetting and booting resources.

Systems composed of multiple but coordinated processors were first developed and used in the context of mainframes. More recently, interest in multiple processor systems has escalated as a consequence of the low cost and high performance of microprocessors, with the objective of replicating mainframe performance through the parallel use of multiple microprocessors.

A variety of architectures have been defined for multi-processor systems. Most designs rely upon highly integrated architectures by virtue of the need for cache coherence. In such systems cache coherence is maintained through complex logic circuit interconnection of the cache memories associated with the individual microprocessors to ensure data consistency as reflected in the various caches and main memory.

A somewhat different approach to architecting a multi-processor system relies upon a relatively loose hardware level coupling of the individual processors, with the singular exception of circuit logic controlling access to the shared global memory, and the use of software to manage cache coherency. An architecture which relies upon software managed cache coherency allows the designer to utilize existing processor hardware to the maximum extent, including the utilization of the processor hardware integrated booting/starting/restarting/resetting resources. This independence of the processors also lends itself to multi-processor systems with accentuated levels of availability, in that such independence facilitates continuity of system operation in the presence of failures or removals of one or more processors. Coordination in the access to, and coherency with, a shared global memory is of course somewhat more difficult with such independence of processors.

A fundamental problem which arises with such individualized processor multi-processor systems involves the coordination to accomplish system wide

booting. Not only are the multiple processors designed and configured to accomplish individualized starting, but such starting must also incorporate the effects of an asynchronous common start signal. The asynchronous signal is usually derived from the status of the power supply. The multi-processor system must also have resources to synchronize the processors undergoing individualized starting to a master clock, and devices and methods to insure initialization and testing of all the processor as well as the shared global memory. Accomplishing this in the face of a failure in one or more of the processors complicates the management of the booting operation, in that booting responsibilities cannot be permanently allocated to selected ones of the processors.

### Summary of the Invention

The present invention provides a method and apparatus for booting a multi-processor computer system having shared global memory and processor local memory, the apparatus comprising: means for asynchronously starting multiple processors; and means for synchronizing to a master clock a signal in the start sequence of each of said multiple processors, wherein the means for synchronising comprises reset circuitry for synchronising a commonly distributed start signal and processor individualised resource start sequences, the reset circuitry being adapted to distribute reset circuit signals to reset both local and global memory.

Preferably, the present invention defines a multiple processor architecture and method of operation in which a plurality of processors having individual starting resources respond to a common start signal and master clock to boot not only their individualized processor resources but the system level global memory. Furthermore, the objectives are attained in the context of an architecture which boots notwithstanding a failure in or absence of one or more of the individualized processors.

In one form, the present invention involves apparatus for booting a multiple processor system having both processor local and shared global memory, wherein the processors have individualized means for starting, the system includes a means for generating a common starting signal to all processors, the system includes a master clock means for synchronizing the multiple processors, and the system includes a means for testing the local and global memory in synchronism with the master clock means and responsive to the common start signal. In another aspect, the invention is directed to methods which perform the steps defined by the apparatus.

A preferred embodiment of the invention involves a multiplicity of processors responsive to individualized off-chip sequencers for processor starting and testing. Each processor has a local memory and ac-

cess to shared global memory through a non-blocking crosspoint switch. Access to global memory is coordinated through an atomic memory access controller, while cache coherence is managed through software. A common start signal is generated in response to the status of a shared power supply and is synchronized to a master clock through reset circuits. The reset circuits also synchronize and coordinate the reset of the local and global memory. Testing of the global memory is accomplished by the first of the processors which reaches a defined state in a program load sequence, which status provides such first processor with access to the global memory while isolating other processors from such access.

The benefits and features of the architecture and methods to which the present invention pertains will be more clearly understood and appreciated upon considering the ensuing description of a detailed embodiment.

#### Brief Description of the Drawings

Figure 1 is a schematic block diagram of a multi-processor system according to an embodiment of the present invention;

Figure 2 is a block diagram showing the circuit functions embodied in a reset circuit according to an embodiment of the present invention;

Figures 3 and 4 illustrate by waveforms the method by which the circuit of Figure 2 operates to accomplish resetting and testing for two different hardware starting conditions.

#### Brief Description of the Preferred Embodiment

Figure 1 illustrates by schematic block diagram an architecture for the multi-processor system to which the present invention pertains. Included within the system are four processors, identified by reference numerals 1-4. A representative example of a processor is the RISC System/6000 workstation with associated AIX Operating System as is commercially available from IBM Corporation. (RISC System/6000 is a registered trademark of International Business Machines Corporation). Each processor includes an off-chip sequencer (OCS) 6 and clock 7, which together cycle the related processor through a sequence of reset and test conditions in anticipation of commencing the initial program load (IPL) to boot the operating system. As conventionally practised, once the off-chip sequencer 6 completes its cycles, the initial program load (IPL) ROM 8 is accessed to commence the booting of the operating system from non-volatile storage such as hard disk (not shown). The multi-processor system in Figure 1 shows the presence of multiple such processors and their individually related starting systems. Associated with each processor 1-4 is a respective and locally addressable

memory block, identified by reference numerals 9, 11, 12 and 13. Though not explicitly shown, each processor also includes a cache type memory for both instructions and data. As noted earlier, cache coherency is managed by software in a manner to be described hereinafter.

The creation of a multi-processor system from a multiplicity of individualized processor systems, including their related starting and memory resources, introduces the need for the other elements. For example, a power good signal on line 14 is distributed to all processor off-chip sequencers to initiate the start sequence. As would be expected, the power good signal on line 14 is asynchronous to the master clock signal on line 16. Therefore, the initiation of the off-chip sequencers and master clock do not coincide. This synchronization problem is further exacerbated by the fact that the embodying off-chip sequencers 6 are often synchronized to their own clocks 7.

Further aspects of the multi-processor system reside in atomic semaphores 17 of atomic controller 15, which allow software to coordinate accesses to the global memory array, generally at 18. The atomic semaphore controller uses lockable semaphore type registers. The atomic semaphore controller will only allow one processor at a time to acquire exclusive access to a semaphore register. However, different processors may own different semaphores at the same time, and each processor may own more than one semaphore at a time. Software uses the semaphores to select which processors can access different blocks of global memory. Software also uses cache flush cycles to maintain global memory coherence with respective processor caches. Atomic counter 19 is for purposes of the multi-processor system boot, used to select the processor which tests global memory array 18 for defects and the like.

Non-blocking cross-point switch 23 uses a relatively conventional design to allow processor 1-4 direct access to the whole of global memory array 18 in the absence of any address contentions. The processors are thereby able to concurrently communicate with the global memory.

The management of the boot operation for the multi-processor system in Figure 1 is accomplished through two reset circuits, reset circuit RC1 at reference numeral 21 and reset circuit RC2 at reference numeral 22. Though a single reset circuit would normally suffice, the embodiment utilizes two because of physical chip size constraints and to minimize timing skews within the system. Reset circuits 21 and 22 are cross-coupled to ensure that both parts of global memory array, namely banks 0-3 and 4-7, are reset at substantially identical times.

Figure 1 also shows the presence of memory isolators 26, interposed between each processor and the memory bus extending to both the local and the

global memory. Memory isolators 26 are used to selectively decouple the off-chip sequencing activities of the three processors which are not performing the global memory test. This avoids extraneous memory bus activity from reaching the global memory while the global memory is being tested by the single selected processor.

The multi-processor boot operation of the embodying system in Figure 1 begins with an asynchronously generated power good signal on line 14. The asynchronous power good signal initiates a multiplicity of asynchronous and individually clocked reset signals in the off-chip sequencers individually associated with each processor. The reset signals emanating from such off-chip sequencers are synchronized to the master clock signal on line 16 in reset circuits 21 and 22, wherein reset circuit 21 accomplishes the reset synchronization for processors 1 and 2 while reset circuit 22 does likewise for processors 3 and 4. The clock synchronized reset signals are conveyed to the respective processors. Also emanating from reset circuits 21 and 22 are clock synchronized reset signals directed to the respective local memories 9, 11, 12 and 13, as well as the respective portions of global memory array 18. In this way, the processors retain substantially independent booting or starting resources yet are synchronized to a common power good type start signal and individualized reset signals using a master clock.

Each off-chip sequencer cycles through multiple states during the course of testing its associated processor. Included within those states are multiple reset cycles which are again synchronized through reset circuits 21 and 22. Each off-chip sequencer concludes with the loading of the initial program load code from ROM 8, which code then initiates the loading of the operating system. As embodied, the initial program load code includes an instruction which directs the processor to read the data in counter 19 of atomic controller 15. Counter 19 is initialized to zero at power up and is incremented after each processor read. Reads by successive processors are serialized, so that no two read the same value. The 0 value identifies to the recipient processor that it is to test not only its own local memory but also the whole of the global memory. In contrast, processors reading non-zero values are directed to test only their respective processor local memories. The requirement that only one processor test the global memory during an interval of time is obviously important, while the selection of the first processor to read the counter draws upon practical considerations. Namely, since the off-chip sequencers are not synchronized, predicting which processor will commence IPL first is not practical. Furthermore, if one or more processors are inoperative, the original design goal requires that the system boot operation still be completed and that only a single processor will undertake to test the global mem-

ory.

In reflection, it should be apparent that this system defines an architecture and related method of operation whereby booting of the system is accomplished without regard to the asynchronous stature of the initiation signal, the asynchronous stature of processor individualized starting sequences, and without regard to the presence or absence of selected processors. For example, if the processor and related resources defined by dashed block 27 in Figure 1 were inoperative, the remaining three processors would boot into a fully operative system in the normal manner.

Figure 2 schematically illustrates the logic internal to an embodying reset circuit. The various blocks are identified by function. The master clock and power good related power on reset (POR) signals are shown together with the inputs and outputs. The source, destination, and character of the input and output signals are defined in the headings of Figures 3 and 4. The timing relationships of the various signals are depicted in Figures 3 and 4. Figure 3 illustrates the waveforms when the hardware reset signal starts at a low level. On the other hand, Figure 4 illustrates the states of the various signals when the hardware reset starts in a high state.

The architecture and method of operation defined by the present invention not only synchronize various asynchronously occurring boot type signals for a multi-processor system having processor local and shared global memory, but accomplishes these objectives with extenuated flexibility. Namely, booting is accomplished with processors having independent starting sequencers, provides master clock synchronized local and global memory reset, and defines a process for selecting a processor to test the global memory. Foremost, these objectives are attainable with one or more processors disconnected.

## Claims

1. Apparatus for booting a multi-processor computer system having shared global memory (18) and processor-local memory (9, 11, 12, 13), comprising:
  - means for asynchronously starting multiple processors (1, 2, 3, 4); and
  - means for synchronizing to a master clock a signal in the start sequence of each of said multiple processors, wherein the means for synchronising comprises reset circuitry (21, 22) for synchronising a commonly distributed start signal and processor individualised resource start sequences, the reset circuitry being adapted to distribute reset circuit signals to reset both local and global memory.

2. Apparatus according to claim 1, further comprising:  
     means for testing (6, 7) global memory by  
     a first processor in synchronism to the master  
     clock. 5
3. Apparatus according to claim 2, wherein said first  
     processor is selected to test the global memory  
     (18) based on said first processor being first  
     among the processors to complete an internal  
     test sequence. 10
4. Apparatus according to claim 3, further comprising:  
     means for decoupling nonselected proces-  
     sors from global memory during the test of global  
     memory. 15
5. Apparatus according to claim 3 or claim 4, where-  
     in the means for testing the global memory using  
     a selected processor further comprises a means  
     for synchronising the global memory reset and  
     the global memory test signals to the master  
     clock means. 20
6. Apparatus according to any one of the preceding  
     claims, wherein the means for asynchronously  
     starting multiple processors includes means for  
     individually starting each processor, the asyn-  
     chronous start sequence in each processor in-  
     volving stepping each processor through a multi-  
     ple stage individually clocked start sequence. 25
7. A method of booting a multi-processor computer  
     system having shared global memory and pro-  
     cessor-local memory, comprising:  
     asynchronously starting the multiple proc-  
     essors (1, 2, 3, 4) and  
     synchronising to a master clock a signal in  
     a start sequence of each of said multiple proces-  
     sors, including synchronising a commonly distrib-  
     uted start signal and processor-individualised re-  
     source start sequences using reset circuitry, re-  
     set signals being distributed to reset both local  
     and global memory. 30
8. A method according to claim 7, further compris-  
     ing the step of:  
     testing the global memory by a first proc-  
     essor in synchronism with the master clock. 35
9. A method according to claim 8, further compris-  
     ing the step of:  
     selecting one of the processors to test the  
     global memory bas d on said one processor be-  
     ing first among the processors to complete an in-  
     ternal test sequence. 40
10. A method according to claim 9, wherein the step  
     of testing the global memory by a selected proc-  
     essor further comprises the synchronisation of  
     the global memory test signals and reset signals  
     to the master clock. 45
11. A method according to any one of claims 7 to 10,  
     wherein the step of asynchronously starting the  
     multiple processors includes:  
     starting each processor individually, the  
     asynchronous start sequence in each processors  
     involving a multiple stage individually clocked  
     start sequence. 50

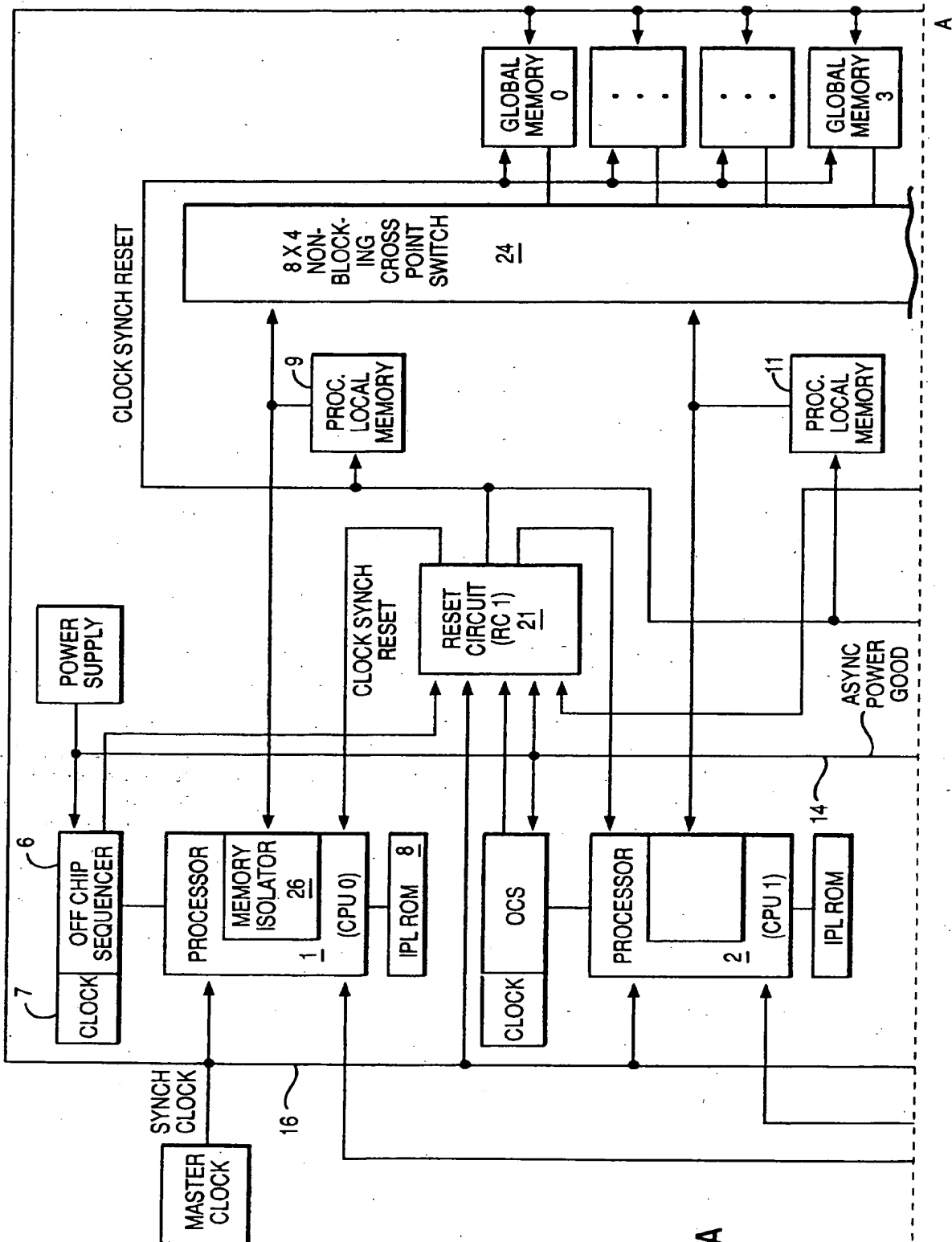


FIG. 1A

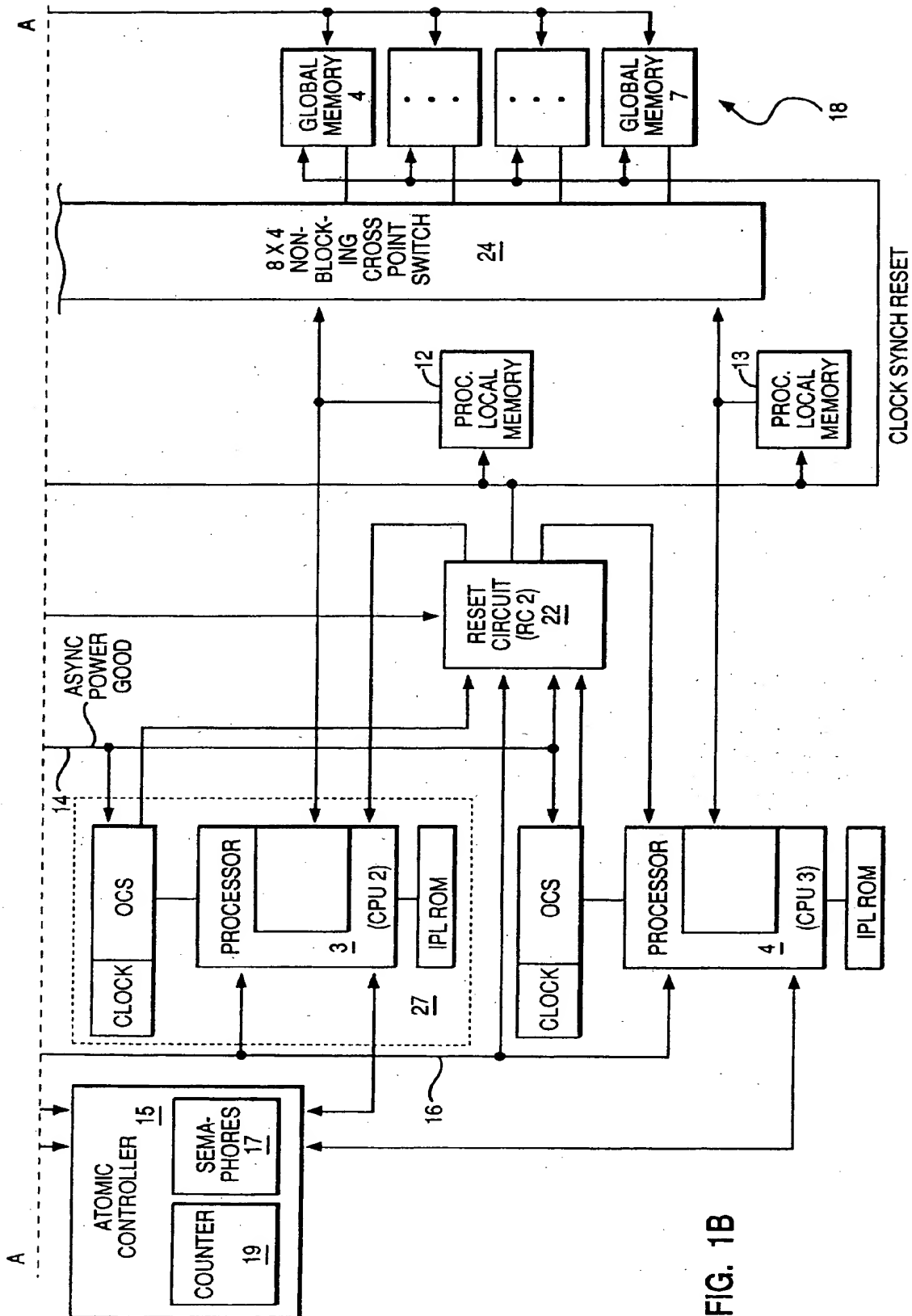


FIG. 1B



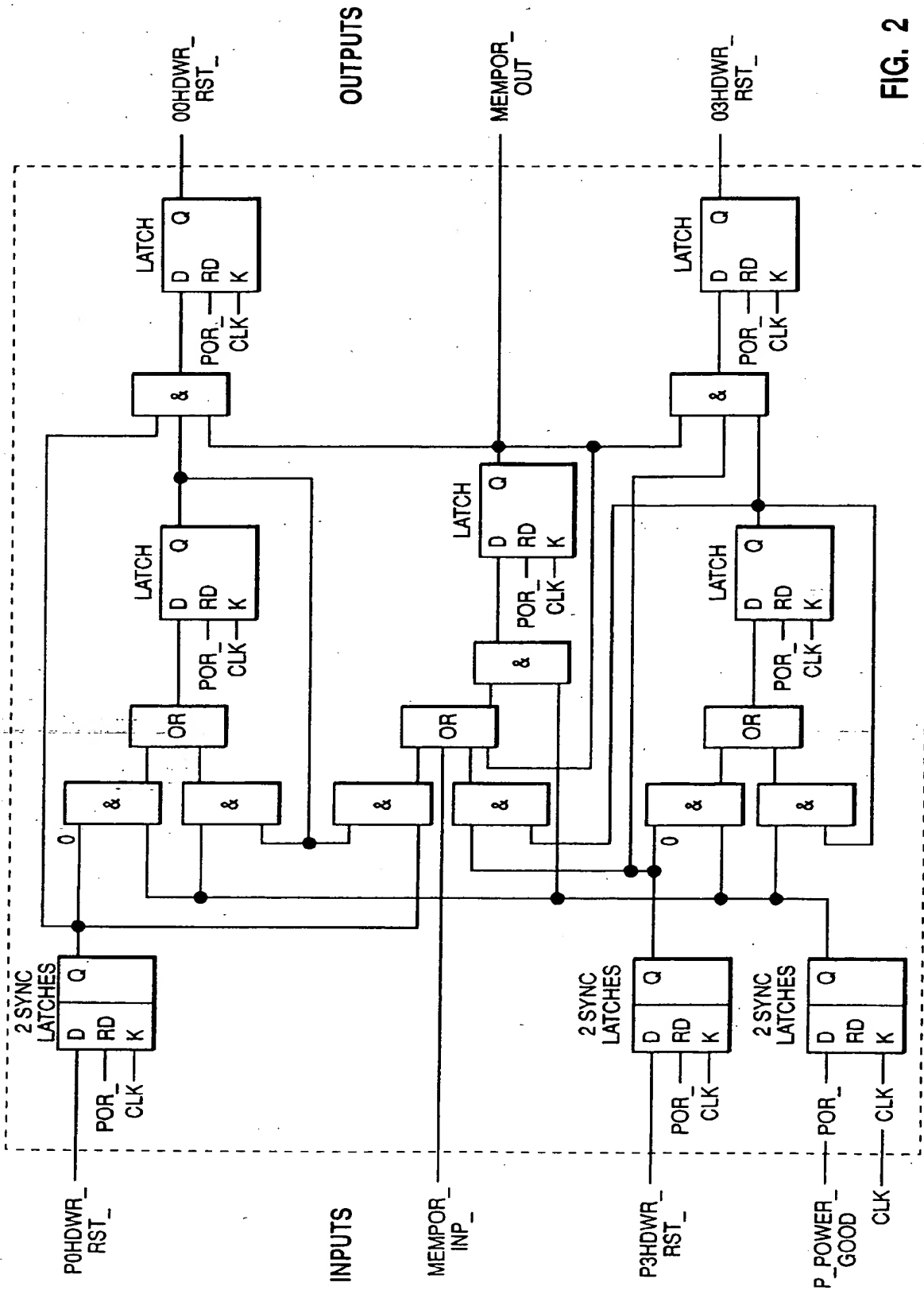
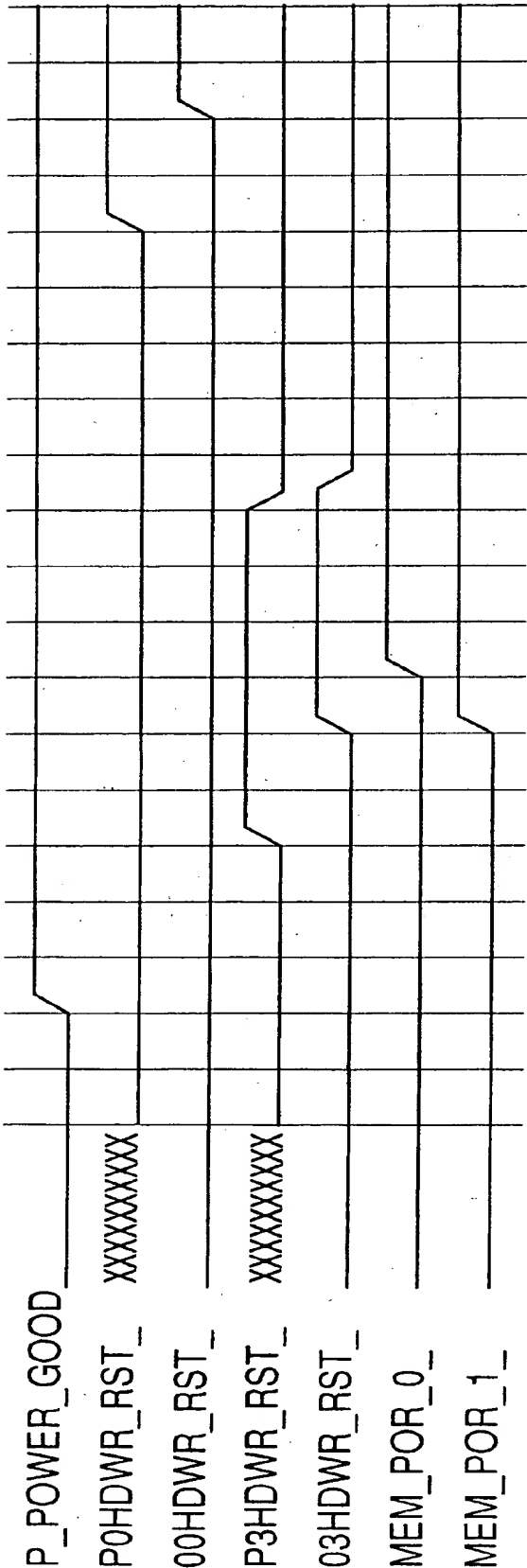


FIG. 2

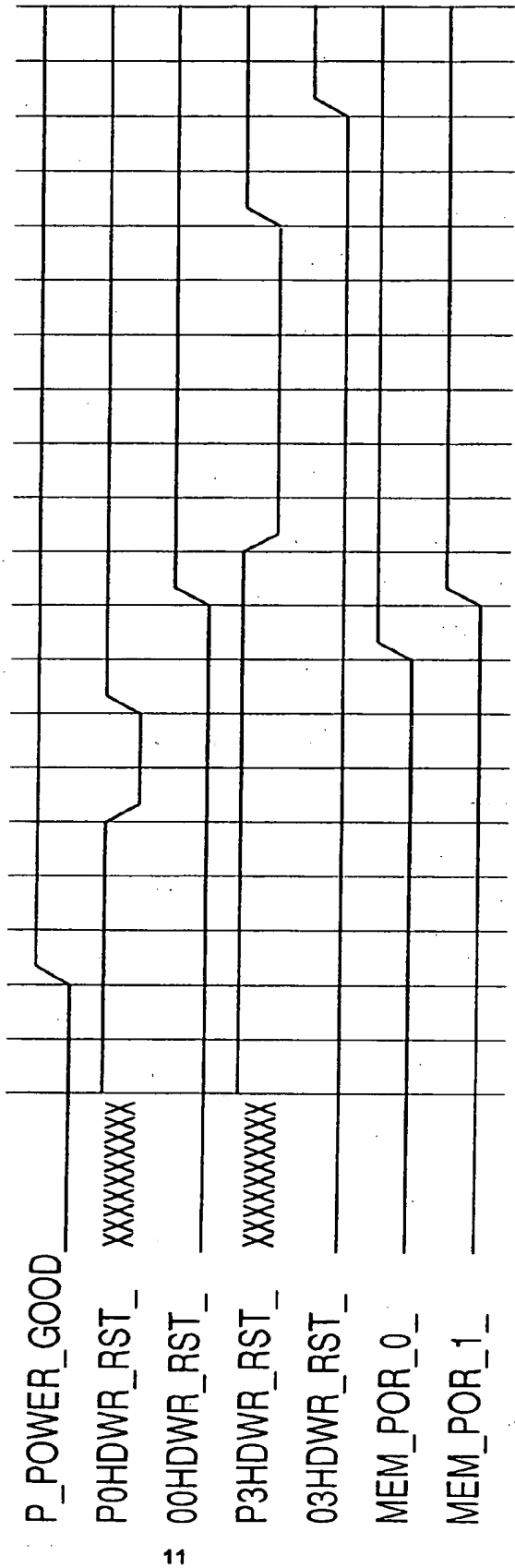
P\_POWER\_GOOD      Signal from the power supply indicating that the power is good. Other Signals should be stable.  
P0HDWR\_RST\_      Asynchronous reset for CPU-0 from OCS-0 (it is synchronized in RC1).  
00HDWR\_RST\_      Synchronous reset for CPU-0 from RC1.  
P3HDWR\_RST\_      Asynchronous reset for CPU-3 from OCS-3 (it is synchronized in RC2).  
03HDWR\_RST\_      Synchronous reset for CPU-3 from RC2.  
MEM\_POR\_0\_      Synchronous reset for Global memory from RC1.  
MEM\_POR\_1\_      Synchronous reset for Global memory from RC2.



HDWR\_RST starting at '0'.

FIG. 3

P\_POWER\_GOOD      Signal from the power supply indicating that the power is good. Other Signals should be stable.  
P0HDWR\_RST\_      Asynchronous reset for CPU-0 from OCS-0 (it is synchronized in RC1).  
00HDWR\_RST\_      Synchronous reset for CPU-0 from RC1.  
P3HDWR\_RST\_      Asynchronous reset for CPU-3 from OCS-3 (it is synchronized in RC2).  
03HDWR\_RST\_      Synchronous reset for CPU-3 from RC2.  
MEM\_POR\_0\_      Synchronous reset for Global memory from RC1.  
MEM\_POR\_1\_      Synchronous reset for Global memory from RC2.



HDWR\_RST starting at '1'.

FIG. 4

**THIS PAGE BLANK (USPTO)**



(11)

**EP 0 602 791 A3**

(12)

**EUROPEAN PATENT APPLICATION**

(88) Date of publication A3:  
19.03.1997 Bulletin 1997/12

(51) Int Cl.<sup>6</sup>: **G06F 15/16**, G06F 9/445

(43) Date of publication A2:  
22.06.1994 Bulletin 1994/25

(21) Application number: 93308706.6

(22) Date of filing: 01.11.1993

(84) Designated Contracting States:  
AT BE CH DE ES FR GB IT LI NL SE

(30) Priority: 30.10.1992 US 969596

(71) Applicant: **International Business Machines Corporation**  
**Armonk, N.Y. 10504 (US)**

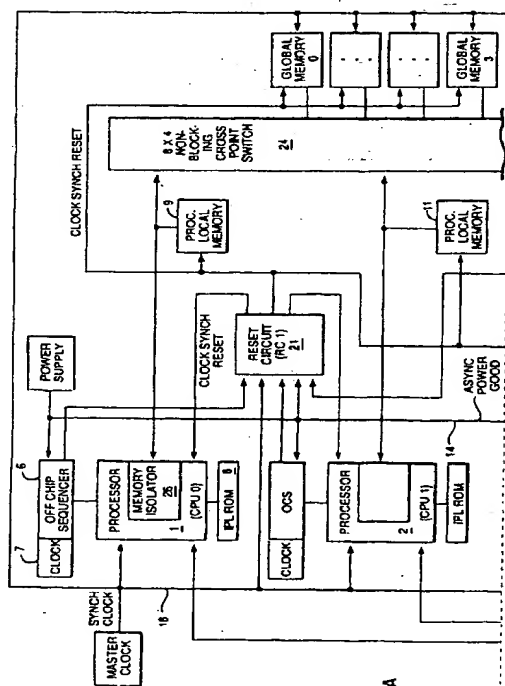
(72) Inventors:  
• **Hardell, William Rudolph, Jr.**  
**Austin, Texas 78759 (US)**

- Henson, James Dodd, Jr.  
Austin, Texas 78758 (US)
- Mitchell, Oscar Reid  
Pflugerville, Texas 78660 (US)

(74) Representative: Moss, Robert Douglas  
IBM United Kingdom Limited  
Intellectual Property Department  
Hursley Park  
Winchester Hampshire SO21 2JN (GB)

(54) **Apparatus and method for booting a multiple processor system having a global/local memory architecture**

(57) An architecture and method for booting a multi-processor system having processor local memory and shared global memory, with shared global memory access managed by an atomic memory access controller and cache coherence managed by software. Reset circuits are used to synchronize to a master clock a commonly distributed start signal and processor individualized restart sequences, which reset circuit signals are distributed to reset both local and global memory. Global memory testing is assigned to a processor based upon its rate status in completing an internal test sequence. The systems and methods are particularly suited to booting a group of multiple but relatively independent processors. Furthermore, the practice of the invention facilitates booting of such system when one or more of the processors have been disconnected or failed.



**FIG. 1A**

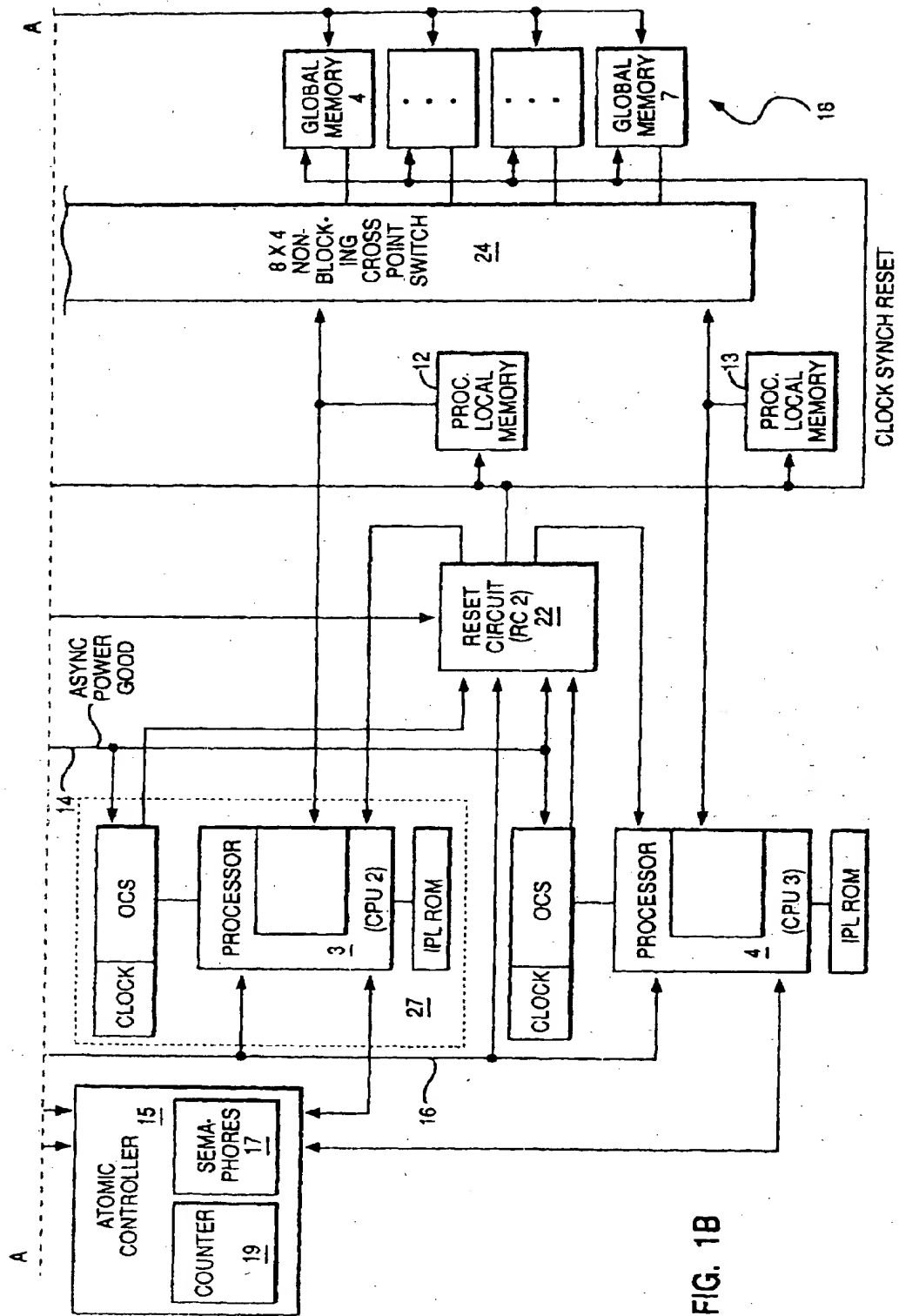


FIG. 1B



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 93 30 8706

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.CLS)
A	EP-A-0 306 348 (DIGITAL EQUIPMENT CORP) 8 March 1989 * page 14, line 47 - page 15, line 4 * * page 19, line 25 - page 20, line 55 * * figure 1 *	1,7	G06F15/16 G06F9/445
A	PATENT ABSTRACTS OF JAPAN vol. 011, no. 144 (P-574), 12 May 1987 & JP-A-61 281352 (HITACHI LTD), 11 December 1986, * abstract *	1,7	
			TECHNICAL FIELDS SEARCHED (Int.CLS)
			G06F
The present search report has been drawn up for all claims			
Place of search BERLIN		Date of completion of the search 8 January 1997	Examiner Masche, C
<p><b>CATEGORY OF CITED DOCUMENTS</b></p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons &amp; : member of the same patent family, corresponding document</p>			

EPO FORM 1503 DIII (P0401)

**THIS PAGE BLANK (USPTO)**